

Exploring School Zones for Family Migrants in the Greater London Area

Rainie Pan

December 14, 2020

1. Introduction

This project tries to identify school zone neighborhoods in the Greater London area of UK and to explore the possible correlations with home price. For new immigrants, one top concern when picking the landing neighbourhood is about having good access to schoolings.

However, as newcomers to the country or even the continent, one may not be familiar with the school districts and neighbourhoods. Further, most new immigrants are often price sensitive and prefer to find school neighbourhoods at a reasonable cost.

We hope to provide the audience with a transparent, data-driven idea about the possible neighbourhoods one can look into.

2. Data Acquisition & Cleansing

To perform the analysis, we collected data from 4 major sources:

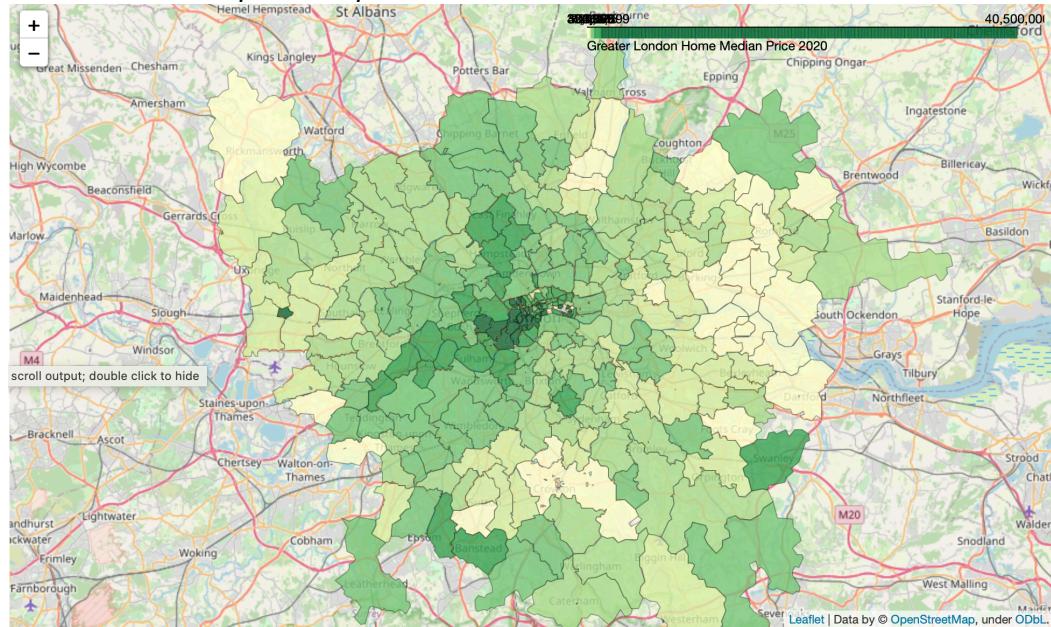
- UK's postcode and geocode data (longitude & latitude) from FreeMapTools ([Link](#))
- UK home price data from Gov.uk ([link](#))
- Greater London area's postal districts and neighbourhoods from Wikipedia
- Venue data from Foursquare API

For the school categories from Foursquare, we selected the following ones:

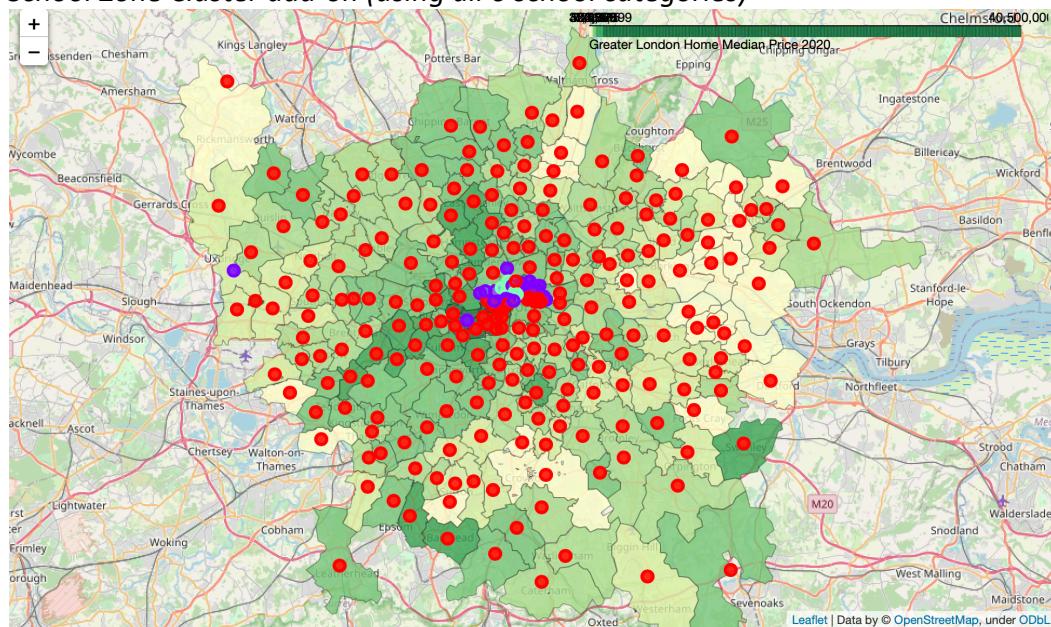
- elementary_ID = '4f4533804b9074f6e4fb0105' #For elementary schools
- nursery_ID = '4f4533814b9074f6e4fb0107' #For nursery schools
- middle_ID = '4f4533814b9074f6e4fb0106' #For middle schools
- high_ID = '4bf58dd8d48988d13d941735' #For high schools
- private_ID = '52e81612bc5c57f1066b7a46' #For private schools
- university_ID = '4bf58dd8d48988d1ae941735' #For university

3. Visualizations

3.1 Home Price Choropleth Map

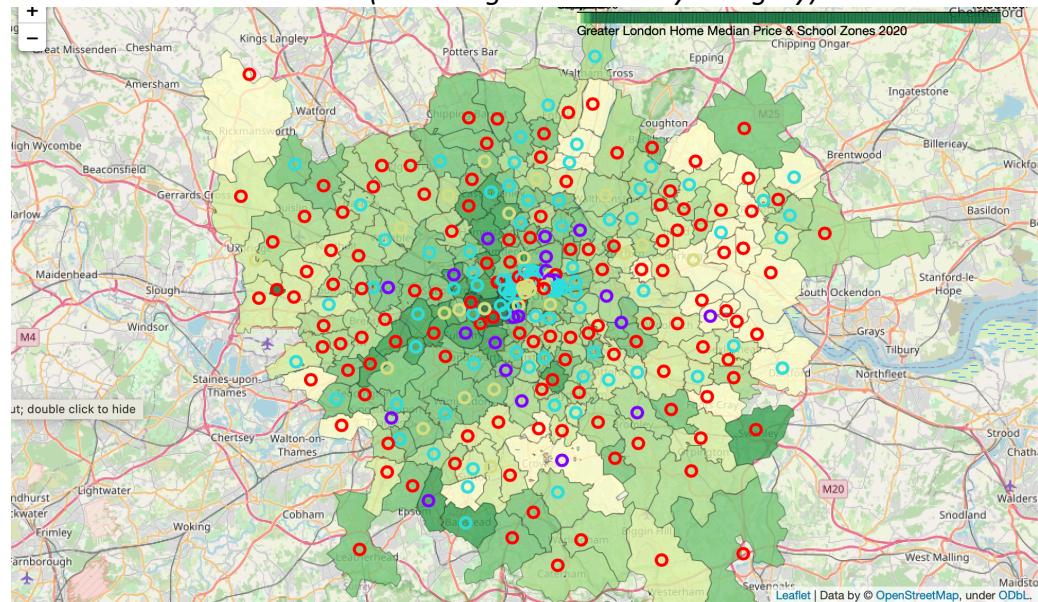


3.2 School Zone Cluster add-on (using all 6 school categories)



In this clustering analysis, we observed clear separation of city center vs. the outer rings, with the exception of one purple dot to the far left. A detailed look at the data revealed that this is due to data skew to university buildings (and the fact that Foursquare API's university venue category returns multiple a large count of university-affiliated buildings). As a result, I decided to drop the university count column for a new clustering analysis.

3.3 School Zone Cluster add-on (excluding the University category)



In this case, we observed more interesting clustering groups and used a final cluster number of 4 based on the elbow test.

4. Preliminary Conclusions

With the analyses above, we have arrived at some preliminary conclusions:

1. For migrant families with kids that require school placement, Cluster 0 (Red) areas generally would have less choices
2. For families with young kids whom go to nursery or elementary schools, Cluster 1 (Purple) areas would have more choices
3. One may consider Cluster 1 areas in the outer rings of Greater London with relatively lower home prices

5. Future Directions

This is a scalable project. The future directions could expand both on the breadth front to cover broader geographies, also on the depth front to obtain more detailed and specific data analytics about the schools and the surrounding neighbourhood.